3.11. Ejercicio de aplicación

En este ejercicio se analizará la estimación de un modelo con y sin ordenada en el origen, su bondad de ajuste y las propiedades de los residuos. También se analizará la estimación de formas funcionales no lineales. Para ello se plantean los siguientes modelos:

$$C_t = \beta_0 + \beta_1 P_t + \beta_2 RF_t + \varepsilon_t \ \forall t = 1, 2, \ldots, 10$$

$$C_t = \beta_1 P_t + \beta_2 RF_t + \varepsilon_t \ \forall t = 1, 2, \ldots, 10$$

En este caso, se utilizarán los 10 primeros datos del fichero aceite.gdt: consumo familiar mensual medio de aceite de oliva en litros (C), precio medio de compra del aceite de oliva en euros/litro (P) e ingresos familiares mensuales medios en euros (RF).

```
? open "C:\Proyecto01\Aceite.gdt"
Leer fichero de datos C:\Proyecto01\Aceite.gdt
periodicidad: 1, máx.obs: 60
rango de observaciones: 1-60

Listando 7 variables:
  0) const     1) C         2) P         3) RF        4) PG
  5) PS        6) TF
```

3.11.1. Estimación MCO del modelo con ordenada en el origen

```
? smpl 1 10
Rango de datos completo: 1 - 60 (n = 60)
Muestra actual: 1 - 10 (n = 10)
```

Obsérvese que aunque se dispone de 60 observaciones para cada variable, sólo se van a utilizar las 10 primeras. Además, el fichero de datos contiene más variables que las que se van a utilizar en la estimación.

```
# Modelo con ordenada en el origen con la opción anova
? ols C const P RF --anova

Modelo 1: MCO, usando las observaciones 1-10
Variable dependiente: C

             Coeficiente   Desv. Típica   Estadístico t   Valor p
  ----------------------------------------------------------------
  const       8.28657      1.28569          6.445       0.0004   ***
  P          -2.93379      0.297456        -9.863       2.34e-05 ***
  RF          0.00907529   0.000854912     10.62        1.44e-05 ***

Media de la vble. dep.  10.30000   D.T. de la vble. dep.   1.337494
Suma de cuad. residuos  0.420867   D.T. de la regresión    0.245202
R-cuadrado              0.973859   R-cuadrado corregido    0.966390
F(2, 7)                 130.3903   Valor p (de F)          2.89e-06
Log-verosimilitud       1.650732   Criterio de Akaike      2.698536
Criterio de Schwarz     3.606292   Crit. de Hannan-Quinn   1.702731


Análisis de Varianza:
                  Suma de cuadrados       gl  Media de cuadrados
  Regresión                15.6791        2             7.83957
  Residuo                   0.420867      7           0.0601239
  Total                    16.1           9             1.78889

  R^2 = 15.6791 / 16.1 = 0.973859
  F(2, 7) = 7.83957 / 0.0601239 = 130.39 [Valor p 2.89e-006]


# Se genera el tamaño muestral
? genr T = $T
Se ha generado el escalar T = 10

# Se genera el número de variables explicativas
? genr k = $ncoeff - 1
Se ha generado el escalar k = 2
```

- Interpretación de resultados

El modelo estimado por MCO es:

$$\hat{C}_t = \underset{(1.28569)}{8.28657} - \underset{(0.297456)}{2.93379}P_t + \underset{(0.000854912)}{0.00907529}RF_t$$

$$t_0 = 6.445 \qquad t_1 = -9.863 \quad t_2 = 10.62$$

$$R^2 = 0.9738$$

Los valores que aparecen entre paréntesis debajo de cada coeficiente estimado (b_i) hacen referencia a su desviación típica estimada $\left(S_{b_i}\right)$ e informan de lo cerca o lejos que está el valor del coeficiente estimado de su valor esperado (informan de su "confiabilidad" o precisión), por ello, cuanto más pequeños sean mejor. El problema que presentan estas desviaciones es que se ven afectadas por las unidades de medida de las variables. Por ello, a veces, entre paréntesis en vez de figurar estas desviaciones figuran las ratios t (t_i), que son adimensionales, de forma que cuanto mayor sea esta ratio en valor absoluto mayor será la precisión del estimador.

De los resultados obtenidos se deduce que:

- El estimador de la ordenada en el origen (8.28657) se podría interpretar como el consumo medio de aceite de oliva de una familia, cuando el nivel de renta mensual y el precio del litro de aceite son nulos. Esta interpretación "mecánica" que se acaba de hacer de b_0 no siempre tiene sentido (este caso es un ejemplo), además cabe señalar que la muestra utilizada no incluye el cero entre los valores observados de dichas variables, por lo que lo más correcto sería interpretar b_0 como el consumo mensual medio de aceite de las familias que no es debido a las variables que aparecen de forma explícita en el modelo.
- Manteniendo constante el nivel de renta, ante un incremento de un euro en el precio del litro de aceite de oliva, la demanda de aceite (medida en términos de consumo mensual por familia), disminuye en 2.93379 litros.
- Manteniendo constante el precio del litro de aceite de oliva, ante un incremento de un euro en el nivel de renta de las familias, la demanda de aceite aumenta en 0.00907529 litros.

Dado que el modelo está formulado con ordenada en el origen, el coeficiente de determinación está acotado entre 0 y 1 y, por tanto, un valor de 0.9738 se puede considerar que representa un elevado poder explicativo (modelo con elevada bondad de ajuste), indicando que aproximadamente el 97% de las variaciones en el consumo mensual de aceite por las familias, son explicadas por las variaciones en el precio del aceite y en el nivel de ingresos familiares mensuales.

3.11.1.1. Resultados de la estimación utilizando álgebra matricial

Una de las ventajas del paquete econométrico Gretl es que permite realizar fácilmente operaciones de cálculo matricial[26]. En este epígrafe, con el fin de que el usuario entienda el procedimiento de estimación y no se quede únicamente en la interpretación de resultados numéricos, se calculan de forma detallada los resultados de la estimación del modelo haciendo uso del álgebra matricial.

- Vector de estimadores

Para obtener los estimadores mínimo cuadráticos ordinarios del vector de parámetros es necesario conocer los elementos de la matriz X'X y del vector X'Y, que se obtendrán a partir de los datos de la muestra:

[26] Las instrucciones para llevar a cabo dichos cálculos ya han sido analizadas en un capítulo anterior.

$$b=(X'X)^{-1}X'Y=\begin{pmatrix} b_0 \\ b_1 \\ b_2 \\ \dots \\ b_K \end{pmatrix}_{(K+1)x1}$$

donde:

- X'X es la matriz de orden (K+1)x(K+1), que contiene los sumatorios de los productos cruzados de los regresores:

$$X'X=\begin{pmatrix} T & \sum_{t=1}^{T} X_{1t} & \dots & \sum_{t=1}^{T} X_{Kt} \\ \sum_{t=1}^{T} X_{1t} & \sum_{t=1}^{T} X_{1t}^2 & \dots & \sum_{t=1}^{T} X_{1t}X_{Kt} \\ \dots & \dots & \ddots & \dots \\ \sum_{t=1}^{T} X_{Kt} & \sum_{t=1}^{T} X_{1t}X_{Kt} & \dots & \sum_{t=1}^{T} X_{Kt}^2 \end{pmatrix}_{(K+1)x(K+1)}$$

- X'Y es el vector de orden (K+1)x1, que contiene los sumatorios de los productos cruzados del regresando con los regresores:

$$X'Y=\begin{pmatrix} \sum_{t=1}^{T} Y_t \\ \sum_{t=1}^{T} Y_t X_{1t} \\ \dots \\ \sum_{t=1}^{T} Y_t X_{Kt} \end{pmatrix}_{(K+1)x1}$$

Recuerde que, siempre que se inicia una sesión de trabajo, Gretl genera de forma automática el regresor ficticio bajo la denominación de **const**.

```
# Se genera la matriz de diseño (X)
? matrix X = {const, P, RF}
Se ha generado la matriz X

# Se muestra la matriz X (que en este caso es de orden 10x3)
? print X
X (10 x 3) [t1 = 1, t2 = 10]
      1.0000       2.4600       1079.5
      1.0000       2.5100       1108.2
      1.0000       2.7300       1116.8
      1.0000       2.7300       978.09
      1.0000       2.6300       809.62
      1.0000       2.3800       1047.1
      1.0000       2.4400       1164.3
      1.0000       2.7500       1069.6
      1.0000       2.9800       1040.3
      1.0000       1.9500       1067.8
```

Cuando se definen matrices a partir de variables, Gretl recuerda el rango muestral con el que se está trabajando, que en este caso comienza en la primera observación (t1=1) y termina en la décima (t2=10).

```
# Se genera el vector que recoge las observaciones del regresando (Y)
? matrix Y = {C}
Se ha generado la matriz Y

# Se muestra el vector Y (en este caso, es de orden 10x1)
? print Y
Y (10 x 1) [t1 = 1, t2 = 10]
  11
  11
  10
   9
   8
  11
  12
  10
   9
  12
```

Nótese que Gretl no ha creado una serie sino un vector columna denominado Y que contiene las 10 primeras observaciones de la serie C. El usuario debe tener esto en cuenta, puesto que determinadas funciones del comando **genr** se pueden utilizar con series pero no con matrices. En el caso de que se desee definir como serie, basta sustituir el comando **genr** por el comando **series** y Gretl añadiría la serie a la lista de variables iniciales.

```
# Se genera la serie que recoge las observaciones del regresando (y)
? series y = {C}
Se ha generado la serie y (ID 7)

# Se muestra la serie y
? print y
Muestra actual:  1 - 10 (n = 10)
11.0000  11.0000  10.0000  9.00000  8.00000  11.0000  12.0000  10.0000
9.00000  12.0000
```

Como puede verse, Gretl asigna a la variable o serie "y", el identificador "ID 7" puesto que en la lista inicial ya tiene 6 variables.

Dado que se ha utilizado la denominación Y para un vector, no se podrá utilizar la misma denominación para una variable, si así se hiciese Gretl emitiría el mensaje de error "*la variable Y es de tipo matrix*":

```
# Se genera la serie que recoge las observaciones del regresando (Y)
? series Y = {C}
La variable Y es de tipo matrix
```

Se debe recordar que se podría utilizar la misma denominación para definir "elementos" de las mismas características, pero que la información nueva sustituiría a la información antigua. Por ejemplo, si se ejecutará el comando:

```
? matrix Y = {P}
Se ha reemplazado la matriz Y

? print Y
Y (10 x 1) [t1 = 1, t2 = 10]
  2.46
  2.51
  2.73
  2.73
  2.63
  2.38
  2.44
  2.75
  2.98
  1.95
```

```
? matrix Y = {C}
Se ha reemplazado la matriz Y

? print Y
Y (10 x 1) [t1 = 1, t2 = 10]
  11
  11
  10
   9
   8
  11
  12
  10
   9
  12
```

Gretl informaria de que el vector Y ha sido reemplazado, dejaría de contener las 10 primeras observaciones de la variable C y pasaría a contener las 10 primeras observaciones de la variable P.

```
# Se calcula y se muestra la matriz de productos cruzados de los regresores
? matrix  XTX = X'* X
Se ha generado la matriz XTX

? print XTX
XTX (3 x 3)
      10.000       25.560       10481.
      25.560       66.038       26743.
      10481.       26743.  1.1072e+007

# Se calcula y se muestra la inversa de X'X
? genr XTXI = inv(XTX)
Se ha generado la matriz XTXI

? print XTXI
XTXI (3 x 3)
      27.493      -4.6269    -0.014852
     -4.6269       1.4716   0.00082567
   -0.014852   0.00082567  1.2156e-005

# Se calcula y se muestra el vector de productos cruzados de los regresores y el regresando
? matrix XTY = X'* Y
Se ha generado la matriz XTY

? print XTY
XTY (3 x 1)
      103.00
      260.76
 1.0888e+005
```

Con la inversa de la matriz X’X y el vector X’Y, se calcula el vector de estimadores MCO:

```
# Se calcula y se muestra el vector de estimadors (b)
? matrix b = XTXI * XTY
Se ha generado la matriz b

? print b
b (3 x 1)
      8.2866
     -2.9338
   0.0090753
```

Cuando se hace un cálculo de este tipo no es necesario realizar todas las operaciones por separado (traspuesta, inversa,...), ya que Gretl permite hacer determinados cálculos y operaciones de una sola vez, por ejemplo, para obtener el vector de estimadores bastaría con una única instrucción:

```
? matrix b = inv(X' * X) * (X' * Y)
```

Además, no sería necesario calcular el vector Y, en este caso, se podría utilizar directamente C en las expresiones de cálculo:

```
? matrix b = inv(X' * X) * (X' * C)
```

- Regresando estimado

$\hat{Y} = \mathrm{Xb}$

```
# Se calcula y se muestra el regresando estimado
? matrix YE = X * b
Se ha generado la matriz YE

? print YE
YE (10 x 1) [t1 = 1, t2 = 10]
      10.866
      10.980
      10.413
      9.1538
      7.9183
      10.807
      11.695
      9.9256
      8.9854
      12.256
```

Si lo que interesa obtener es la estimación para una única observación del regresando, se puede utilizar la expresión:

$\hat{Y}_t = X_t'\mathrm{b} = \mathrm{b}_0 + \mathrm{b}_1 X_{1t} + \cdots + \mathrm{b}_K X_{Kt}$

```
# Se calcula el valor estimado para la primera observación
? genr YE1 = b[1] + b[2] * X[1,2] + b[3] * X[1,3]
Se ha generado el escalar YE1 = 10.8663
```

Recuerde que cuando se utiliza el comando **genr** y el resultado es un escalar, Gretl informa de su valor y no será necesario utilizar un comando **print** para visualizarlo (sería necesario en el caso de generar series o matrices).

Los valores de los regresores para una observación concreta se pueden obtener a partir de la matriz de diseño (X) o a partir de las series de datos, que son reconocidas por Gretl como vectores columna[27]:

```
? matrix YE1M = b[1] + b[2] * P[1] + b[3] * RF[1]
Se ha generado la matriz YE1M

? print YE1M
YE1M (1 x 1)
      10,866
```

Para generar la serie que contiene los valores estimados del regresando:

```
# Se genera y se muestra la serie que recoge las observaciones del regresando estimado (ye)
? series ye = X * b
Se ha generado la serie ye (ID 8)

? print ye
Muestra actual:  1 - 10 (n = 10)
10.8663  10.9802  10.4130  9.15379  7.91825  10.8067  11.6949  9.92559
8.98537  12.2559
```

- Residuos

$e = Y - \hat{Y}$

```
# Se calcula y se muestra el vector de residuos
? matrix E = Y - YE
Se ha generado la matriz E

? print E
E (10 x 1) [t1 = 1, t2 = 10]
```

[27] Recuerde que esta afirmación no se cumple a la inversa, es decir, los vectores columna no son reconocidos por Gretl como series.

```
     0.13367
    0.019811
    -0.41298
    -0.15379
    0.081748
     0.19328
     0.30514
    0.074408
    0.014631
    -0.25592
```

Para obtener un residuo para una observación concreta:

$e_t = Y_t - \hat{Y}_t$

```
# Se calcula el valor estimado para la primera observación
? genr e1 = Y[1] - YE[1]
Se ha generado el escalar e1 = 0.133674
```

Para obtener la serie de residuos:

```
# Se genera y se muestra la serie de residuos
? series e = Y - YE
Se ha generado la serie e (ID 9)

? print e
Muestra actual:  1 - 10 (n = 10)
 0.133674   0.0198114  -0.412985  -0.153788   0.0817480   0.193283
 0.305141   0.0744077   0.0146306  -0.255922
```

- Suma de Cuadrados Totales

$SCT = \sum_{t=1}^{T}(Y_t - \bar{Y})^2 = Y'Y - T\,\bar{Y}^2$

```
# Se calcula la SCT
? genr SCT = sum((y - mean(y))^2)
Se ha generado el escalar SCT = 16.1

# Otra alternativa de cálculo
? genr SCT = Y' * Y - T * (mean(Y))^2
Se ha reemplazado el escalar SCT = 16.1
```

- Suma de Cuadrados de Regresión

$SCR = \sum_{t=1}^{T}\left(\hat{Y}_t - \bar{\hat{Y}}\right)^2 = \hat{Y}'\hat{Y} - T\,\bar{\hat{Y}}^2 = b'X'Y - T\,\bar{Y}^2$

```
# Se calcula la SCR
? genr SCR = sum((ye - mean(ye))^2)
Se ha generado el escalar SCR = 15.6791

# Otras alternativas de cálculo
? genr SCR = YE' * YE - T * (mean(YE))^2
Se ha reemplazado el escalar SCR = 15.6791

? genr SCR = b' * X' * Y - T * (mean(YE))^2
Se ha reemplazado el escalar SCR = 15.6791
```

- Suma de Cuadrados de Errores

$SCE = e'e = \sum_{t=1}^{T} e_t^2 = Y'Y - b'X'Y$

```
# Se calcula la SCE
? genr SCE = sum(e^2)
Se ha generado el escalar SCE = 0.420867

# Otras alternativas de cálculo
? genr SCE = Y' * Y - b' * X' * Y
Se ha reemplazado el escalar SCE = 0.420867

? genr SCE = E' * E
Se ha reemplazado el escalar SCE = 0.420867
```

Al estar el modelo formulado con ordenada en el origen, también se podría calcular SCE como diferencia entre SCT y SCR:

```
? genr SCE = SCT - SCR
```

- Estimador de la varianza de la perturbación

$$S^2 = \frac{SCE}{T - K - 1}$$

```
# Se calcula el estimador de la varianza de la perturbación
? genr s2 = SCE / (T - k -1)
Se ha generado el escalar s2 = 0.0601239

# Se calcula el estimador de la desviación típica de la perturbación
? genr s = sqrt(s2)
Se ha generado el escalar s = 0.245202
```

- Matriz de varianzas-covarianzas estimada de los EMCO

$$\hat{V}(b) = S^2 (X'X)^{-1} = \begin{pmatrix} S^2_{b_0} & \cdots & \cdots & \cdots & \cdots \\ S_{b_1 b_0} & S^2_{b_1} & \cdots & \cdots & \cdots \\ S_{b_2 b_0} & S_{b_2 b_1} & S^2_{b_2} & \cdots & \cdots \\ \cdots & \cdots & \cdots & \ddots & \cdots \\ S_{b_K b_0} & S_{b_K b_1} & S_{b_K b_2} & \cdots & S^2_{b_K} \end{pmatrix}_{(K+1)x(K+1)}$$

```
# Se calcula y se muestra la matriz de varianzas-covarianzas estimada de los estimadores
? matrix Vb = s2 * inv(X' * X)
Se ha generado la matriz Vb

? print Vb
Vb (3 x 3)
      1.6530     -0.27819  -0.00089295
    -0.27819     0.088480  4.9642e-005
 -0.00089295  4.9642e-005  7.3088e-007

# Se calcula y se muestra el vector que contiene las varianzas estimadas de los EMCO
? matrix s2b = diag(Vb)
Se ha generado la matriz s2b

? print s2b
s2b (3 x 1)
      1.6530
    0.088480
 7.3088e-007

# Se calcula y se muestra el vector que contiene las desviaciones típicas estimadas de los EMCO
? matrix sb = sqrt(s2b)
Se ha generado la matriz sb

? print sb
sb (3 x 1)
      1.2857
     0.29746
  0.00085491
```

- Ratios t de los EMCO

$$t_i = \frac{b_i}{S_{b_i}}$$

```
# Se calcula las ratios t
? genr t0 = b[1] / sb[1]
Se ha generado el escalar t0 = 6.44524

? genr t1 = b[2] / sb[2]
Se ha generado el escalar t1 = -9.86292

? genr t2 = b[3] / sb[3]
```

```
Se ha generado el escalar t2 = 10.6155
```

Otra forma de calcular los ratios t sería a través del comando **matrix** y de su operador **división elemento a elemento**:

```
? matrix t = b ./ sb
Se ha generado la matriz t

? print t
t (3 x 1)
      6,4452
     -9,8629
      10,615
```

Nótese que en este caso se generaría un vector columna de orden 3x1.

- Coeficiente de determinación y Tabla ANOVA respecto a la media

$$R^2 = \frac{SCR}{SCT}$$

```
# Cálculo del coeficiente de determinación
? genr R2 = SCR/SCT
Se ha generado el escalar R2 = 0.973859
```

Como el modelo está formulado con ordenada en el origen, aplicando la descomposición de SCT en SCE y SCR, el R^2 también se puede obtener como:

$$R^2 = 1 - \frac{SCE}{SCT}$$

```
? genr R2 = 1 - SCE/SCT
```

La información necesaria para el cálculo del coeficiente de determinación en modelos formulados con ordenada en el origen está disponible en la tabla de **Análisis de Varianza**[28] que proporciona Gretl al ejecutar un comando **ols** con la opción **--anova**:

```
Análisis de Varianza:

                  Suma de cuadrados       gl  Media de cuadrados

  Regresión                15.6791        2            7.83957
  Residuo                   0.420867      7          0.0601239
  Total                    16.1           9            1.78889

  R^2 = 15.6791 / 16.1 = 0.973859
  F(2, 7) = 7.83957 / 0.0601239 = 130.39 [Valor p 2.89e-006]
```

- Tabla ANOVA respecto al origen y coeficiente de determinación bruto

La descomposición de SCT en SCE y SCR, sólo se cumple en modelos formulados con ordenada en el origen. Pero la descomposición de $\sum_{t=1}^{T} Y_t^2$ en $\sum_{t=1}^{T} \hat{Y}_t^2$ y $\sum_{t=1}^{T} e_t^2 = SCE$ se cumple siempre (modelos formulados con o sin ordenada en el origen[29]).

Aunque no aparezca entre las salidas de Gretl, parece conveniente calcular la tabla ANOVA respecto al origen, por ser la base para calcular el coeficiente de determinación en modelos formulados sin ordenada en el origen (coeficiente de determinación bruto):

```
# ANOVA respecto al origen
? genr scy = sum(y^2)
Se ha generado el escalar scy = 1077

? genr scye = sum(ye^2)
```

[28] Otros paquetes econométricos como Shazam, la denominan ANOVA respecto a la media.

[29] Algunos paquetes econométricos como Shazam, por analogía recogen esta descomposición en una tabla denominada, ANOVA respecto al origen.

```
Se ha generado el escalar scye = 1076.58

? genr sce = sum(e^2)
Se ha generado el escalar sce = 0.420867
```

Nótese que la única diferencia con la tabla ANOVA respecto a la media, es que en la tabla ANOVA respecto al origen, la suma de cuadrados de las desviaciones no se calculan respecto a la media sino respecto al cero. Puede observarse que la suma de cuadrados de los residuos respecto a su media y respecto al cero coinciden puesto que, en modelos formulados con ordenada en el origen, la media de los residuos es cero (véase **Tabla 3-2**).

Gretl sólo proporciona la tabla ANOVA respecto a la media cuando el modelo está formulado con ordenada en el origen, en caso contrario, no muestra ninguna tabla de análisis de la varianza.

Tabla 3-2. Tabla **ANOVA** respecto al origen.

Análisis de Varianza:

		Sumas de cuadrados	gl	Media de cuadrados
Regresión	$\sum_{t=1}^{T}(\hat{Y}_t-0)^2 = \sum_{t=1}^{T}\hat{Y}_t^2$	1076.58	3	358.86
Residuo	$\sum_{t=1}^{T}(e_t-0)^2 = \sum_{t=1}^{T}e_t^2 = SCE$	0.420867	7	0.0601239
Total	$\sum_{t=1}^{T}(Y_t-0)^2 = \sum_{t=1}^{T}Y_t^2$	1077	10	107.7

R^2 = 1076.58 / 1077 = 0.9996

F(3, 7) = 358.86 / 0.0601239 = 5968.674 [Valor p 2.74605e-012]

Observe que la SCE puede calcularse utilizando ambas tablas.

```
# Cálculo del coeficiente de determinación bruto
? genr R2RAW = scye/scy
Se ha generado el escalar R2RAW = 0.999609

? genr R2RAW = 1 - sce/scy
Se ha reemplazado el escalar R2RAW = 0.999609
```

- Coeficiente de determinación corregido o ajustado

$$\bar{R}^2 = 1 - \frac{\frac{SCE}{T-K-1}}{\frac{SCT}{T-1}}$$

```
# Se calcula el coeficiente de determinación ajustado
? genr R2A = 1 - ((SCE / (T - k - 1)) / (SCT / (T - 1)))
Se ha generado el escalar R2A = 0.96639
```

- Matrices H y M

La matriz H se define como:

$$H = X(X'X)^{-1}X'$$

```
# Se calcula la matriz H
? matrix H = X * inv(X' * X) * X'
Se ha generado la matriz H
? print H
H (10 x 10) [t1 = 1, t2 = 10]
0.12055      0.12345      0.10068     0.058765     0.019414      0.11999      0.14848     0.084099     0.048727      0.17584
0.12345      0.14243      0.14442     0.048353    -0.066478      0.10245      0.18255      0.11135     0.086936      0.12455
0.10068      0.14442      0.22167     0.085860     -0.11031     0.043911      0.17745      0.18168      0.22500    -0.070346
0.058765     0.048353     0.085860      0.18408      0.28351     0.065864   -0.0052496      0.12327      0.18956    -0.034012
0.019414    -0.066478     -0.11031      0.28351      0.77050      0.11851     -0.21961     0.022011     0.084786     0.097665
0.11999      0.10245     0.043911     0.065864      0.11851      0.14591      0.11176     0.046188   -0.0089576      0.25438
0.14848      0.18255      0.17745   -0.0052496     -0.21961      0.11176      0.26168      0.11375     0.058040      0.17114
0.084099      0.11135      0.18168      0.12327     0.022011     0.046188      0.11375      0.16786      0.22528    -0.075487
```

```
0.048727     0.086936      0.22500      0.18956     0.084786   -0.0089576     0.058040      0.22528      0.35985     -0.26922
0.17584      0.12455     -0.070346    -0.034012    0.097665      0.25438      0.17114    -0.075487     -0.26922      0.62548
```

H se denomina **matriz de proyección**, pues proyecta la variable dependiente sobre el espacio de las variables independientes, obteniendo el regresando estimado:

$\hat{Y} = HY$

```
matrix HY=H*Y
Se ha generado la matriz HY

? print HY YE
HY (10 x 1) [t1 = 1, t2 = 10]
      10.866
      10.980
      10.413
      9.1538
      7.9183
      10.807
      11.695
      9.9256
      8.9854
      12.256

YE (10 x 1) [t1 = 1, t2 = 10]
      10.866
      10.980
      10.413
      9.1538
      7.9183
      10.807
      11.695
      9.9256
      8.9854
      12.256
```

La matriz M se define como:

$M = I - H$

```
# Se calcula la matriz M
? matrix M = I(T) - H
Se ha generado la matriz M
? print M
M (10 x 10) [t1 = 1, t2 = 10]
0.87945     -0.12345     -0.10068    -0.058765    -0.019414     -0.11999     -0.14848    -0.084099    -0.048727     -0.17584
-0.12345      0.85757     -0.14442    -0.048353     0.066478     -0.10245     -0.18255     -0.11135    -0.086936     -0.12455
-0.10068     -0.14442      0.77833    -0.085860      0.11031    -0.043911     -0.17745     -0.18168     -0.22500     0.070346
-0.058765    -0.048353    -0.085860      0.81592     -0.28351    -0.065864    0.0052496     -0.12327     -0.18956     0.034012
-0.019414     0.066478      0.11031     -0.28351      0.22950     -0.11851      0.21961    -0.022011    -0.084786    -0.097665
-0.11999     -0.10245    -0.043911    -0.065864     -0.11851      0.85409     -0.11176    -0.046188    0.0089576     -0.25438
-0.14848     -0.18255     -0.17745    0.0052496      0.21961     -0.11176      0.73832     -0.11375    -0.058040     -0.17114
-0.084099     -0.11135     -0.18168     -0.12327    -0.022011    -0.046188     -0.11375      0.83214     -0.22528     0.075487
-0.048727    -0.086936     -0.22500     -0.18956    -0.084786    0.0089576    -0.058040     -0.22528      0.64015      0.26922
-0.17584     -0.12455     0.070346     0.034012    -0.097665     -0.25438     -0.17114     0.075487      0.26922      0.37452
```

M proyecta la variable dependiente sobre el espacio ortogonal a ella o espacio residual, obteniendo el error de estimación:

$e = MY$

```
? matrix MY = M * Y
Se ha generado la matriz MY

? print MY E
MY (10 x 1) [t1 = 1, t2 = 10]
     0.13367
    0.019811
    -0.41298
    -0.15379
    0.081748
     0.19328
     0.30514
    0.074408
    0.014631
```

```
    -0.25592

E (10 x 1) [t1 = 1, t2 = 10]
     0.13367
    0.019811
    -0.41298
    -0.15379
    0.081748
     0.19328
     0.30514
    0.074408
    0.014631
    -0.25592
```

Se comprueba fácilmente que son matrices simétricas, idempotentes y ortogonales entre sí:

```
# Simetría de H: H' = H
? matrix HT = H'
Se ha generado la matriz HT

? print  HT
HT (10 x 10)
0.12055      0.12345      0.10068     0.058765     0.019414      0.11999      0.14848     0.084099     0.048727      0.17584
0.12345      0.14243      0.14442     0.048353    -0.066478      0.10245      0.18255      0.11135     0.086936      0.12455
0.10068      0.14442      0.22167     0.085860     -0.11031     0.043911      0.17745      0.18168      0.22500    -0.070346
0.058765     0.048353     0.085860      0.18408      0.28351     0.065864   -0.0052496      0.12327      0.18956    -0.034012
0.019414    -0.066478     -0.11031      0.28351      0.77050      0.11851     -0.21961     0.022011     0.084786     0.097665
0.11999      0.10245     0.043911     0.065864      0.11851      0.14591      0.11176     0.046188   -0.0089576      0.25438
0.14848      0.18255      0.17745   -0.0052496     -0.21961      0.11176      0.26168      0.11375     0.058040      0.17114
0.084099      0.11135      0.18168      0.12327     0.022011     0.046188      0.11375      0.16786      0.22528    -0.075487
0.048727     0.086936      0.22500      0.18956     0.084786   -0.0089576     0.058040      0.22528      0.35985     -0.26922
0.17584      0.12455    -0.070346    -0.034012     0.097665      0.25438      0.17114    -0.075487     -0.26922      0.62548

# Idempotencia de H: HH = HHH = ... = H
? matrix HH = H * H
Se ha generado la matriz HH

? print HH
HH (10 x 10) [t1 = 1, t2 = 10]
0.12055      0.12345      0.10068     0.058765     0.019414      0.11999      0.14848     0.084099     0.048727      0.17584
0.12345      0.14243      0.14442     0.048353    -0.066478      0.10245      0.18255      0.11135     0.086936      0.12455
0.10068      0.14442      0.22167     0.085860     -0.11031     0.043911      0.17745      0.18168      0.22500    -0.070346
0.058765     0.048353     0.085860      0.18408      0.28351     0.065864   -0.0052496      0.12327      0.18956    -0.034012
0.019414    -0.066478     -0.11031      0.28351      0.77050      0.11851     -0.21961     0.022011     0.084786     0.097665
0.11999      0.10245     0.043911     0.065864      0.11851      0.14591      0.11176     0.046188   -0.0089576      0.25438
0.14848      0.18255      0.17745   -0.0052496     -0.21961      0.11176      0.26168      0.11375     0.058040      0.17114
0.084099      0.11135      0.18168      0.12327     0.022011     0.046188      0.11375      0.16786      0.22528    -0.075487
0.048727     0.086936      0.22500      0.18956     0.084786   -0.0089576     0.058040      0.22528      0.35985     -0.26922
0.17584      0.12455    -0.070346    -0.034012     0.097665      0.25438      0.17114    -0.075487     -0.26922      0.62548

# Por simetría e idempotencia de H: H'H = HH' = H
? matrix HTH = H'* H
Se ha generado la matriz HTH

? print HTH
HTH (10 x 10) [t1 = 1, t2 = 10]
0.12055      0.12345      0.10068     0.058765     0.019414      0.11999      0.14848     0.084099     0.048727      0.17584
0.12345      0.14243      0.14442     0.048353    -0.066478      0.10245      0.18255      0.11135     0.086936      0.12455
0.10068      0.14442      0.22167     0.085860     -0.11031     0.043911      0.17745      0.18168      0.22500    -0.070346
0.058765     0.048353     0.085860      0.18408      0.28351     0.065864   -0.0052496      0.12327      0.18956    -0.034012
0.019414    -0.066478     -0.11031      0.28351      0.77050      0.11851     -0.21961     0.022011     0.084786     0.097665
0.11999      0.10245     0.043911     0.065864      0.11851      0.14591      0.11176     0.046188   -0.0089576      0.25438
0.14848      0.18255      0.17745   -0.0052496     -0.21961      0.11176      0.26168      0.11375     0.058040      0.17114
0.084099      0.11135      0.18168      0.12327     0.022011     0.046188      0.11375      0.16786      0.22528    -0.075487
0.048727     0.086936      0.22500      0.18956     0.084786   -0.0089576     0.058040      0.22528      0.35985     -0.26922
0.17584      0.12455    -0.070346    -0.034012     0.097665      0.25438      0.17114    -0.075487     -0.26922      0.62548

# Simetría de M: M' = M
? matrix MT = M'
Se ha generado la matriz MT

? print MT
MT (10 x 10)
0.87945     -0.12345     -0.10068    -0.058765    -0.019414     -0.11999     -0.14848    -0.084099    -0.048727     -0.17584
-0.12345      0.85757     -0.14442    -0.048353     0.066478     -0.10245     -0.18255     -0.11135    -0.086936     -0.12455
-0.10068     -0.14442      0.77833    -0.085860      0.11031    -0.043911     -0.17745     -0.18168     -0.22500     0.070346
-0.058765    -0.048353    -0.085860      0.81592     -0.28351    -0.065864    0.0052496     -0.12327     -0.18956     0.034012
-0.019414     0.066478      0.11031     -0.28351      0.22950     -0.11851      0.21961    -0.022011    -0.084786    -0.097665
-0.11999     -0.10245    -0.043911    -0.065864     -0.11851      0.85409     -0.11176    -0.046188    0.0089576     -0.25438
-0.14848     -0.18255     -0.17745    0.0052496      0.21961     -0.11176      0.73832     -0.11375    -0.058040     -0.17114
-0.084099     -0.11135     -0.18168     -0.12327    -0.022011    -0.046188     -0.11375      0.83214     -0.22528     0.075487
-0.048727    -0.086936     -0.22500     -0.18956    -0.084786    0.0089576    -0.058040     -0.22528      0.64015      0.26922
-0.17584     -0.12455     0.070346     0.034012    -0.097665     -0.25438     -0.17114     0.075487      0.26922      0.37452
```

```
# Idempotencia de M: MM = MMM = ... = M
? matrix MM = M * M
Se ha generado la matriz MM

? print MM
MM (10 x 10) [t1 = 1, t2 = 10]
0.87945     -0.12345     -0.10068     -0.058765    -0.019414    -0.11999     -0.14848     -0.084099    -0.048727    -0.17584
-0.12345     0.85757     -0.14442     -0.048353     0.066478    -0.10245     -0.18255     -0.11135     -0.086936    -0.12455
-0.10068    -0.14442      0.77833     -0.085860     0.11031     -0.043911    -0.17745     -0.18168     -0.22500      0.070346
-0.058765   -0.048353    -0.085860     0.81592     -0.28351     -0.065864     0.0052496   -0.12327     -0.18956      0.034012
-0.019414    0.066478     0.11031     -0.28351      0.22950     -0.11851      0.21961     -0.022011    -0.084786    -0.097665
-0.11999    -0.10245     -0.043911    -0.065864    -0.11851      0.85409     -0.11176     -0.046188     0.0089576   -0.25438
-0.14848    -0.18255     -0.17745      0.0052496    0.21961     -0.11176      0.73832     -0.11375     -0.058040    -0.17114
-0.084099   -0.11135     -0.18168     -0.12327     -0.022011    -0.046188    -0.11375      0.83214     -0.22528      0.075487
-0.048727   -0.086936    -0.22500     -0.18956     -0.084786     0.0089576   -0.058040    -0.22528      0.64015      0.26922
-0.17584    -0.12455      0.070346     0.034012    -0.097665    -0.25438     -0.17114      0.075487     0.26922      0.37452

# Por simetría e idempotencia de M: M'M = MM' = M
? matrix MTM = M'* M
Se ha generado la matriz MTM

? print MTM
MTM (10 x 10) [t1 = 1, t2 = 10]
0.87945     -0.12345     -0.10068     -0.058765    -0.019414    -0.11999     -0.14848     -0.084099    -0.048727    -0.17584
-0.12345     0.85757     -0.14442     -0.048353     0.066478    -0.10245     -0.18255     -0.11135     -0.086936    -0.12455
-0.10068    -0.14442      0.77833     -0.085860     0.11031     -0.043911    -0.17745     -0.18168     -0.22500      0.070346
-0.058765   -0.048353    -0.085860     0.81592     -0.28351     -0.065864     0.0052496   -0.12327     -0.18956      0.034012
-0.019414    0.066478     0.11031     -0.28351      0.22950     -0.11851      0.21961     -0.022011    -0.084786    -0.097665
-0.11999    -0.10245     -0.043911    -0.065864    -0.11851      0.85409     -0.11176     -0.046188     0.0089576   -0.25438
-0.14848    -0.18255     -0.17745      0.0052496    0.21961     -0.11176      0.73832     -0.11375     -0.058040    -0.17114
-0.084099   -0.11135     -0.18168     -0.12327     -0.022011    -0.046188    -0.11375      0.83214     -0.22528      0.075487
-0.048727   -0.086936    -0.22500     -0.18956     -0.084786     0.0089576   -0.058040    -0.22528      0.64015      0.26922
-0.17584    -0.12455      0.070346     0.034012    -0.097665    -0.25438     -0.17114      0.075487     0.26922      0.37452

# Ortogonalidad ente H y M: HM = 0
? matrix HM=H*M
Se ha generado la matriz HM

? print HM
HM (10 x 10) [t1 = 1, t2 = 10]
-3.3168e-015 -3.4367e-015 -3.7651e-015 -3.5082e-015 -3.0315e-015 -3.1397e-015 -3.4553e-015 -3.7345e-015 -4.0288e-015 -2.5465e-015
-3.3956e-015 -3.6187e-015 -3.6891e-015 -3.1823e-015 -2.4960e-015 -3.2562e-015 -3.7051e-015 -3.5686e-015 -3.5450e-015 -3.1941e-015
-3.8027e-015 -3.7375e-015 -3.0465e-015 -2.7001e-015 -2.5992e-015 -4.0509e-015 -4.1434e-015 -2.9497e-015 -2.0302e-015 -5.3330e-015
-3.6041e-015 -3.2643e-015 -2.7506e-015 -3.8676e-015 -5.4187e-015 -4.1121e-015 -3.0234e-015 -3.0247e-015 -2.8879e-015 -4.7282e-015
-3.2263e-015 -2.6346e-015 -2.7030e-015 -5.4900e-015 -8.7276e-015 -3.7391e-015 -1.4166e-015 -3.6786e-015 -4.3837e-015 -3.0334e-015
-3.1035e-015 -3.3144e-015 -3.9323e-015 -3.9056e-015 -3.6173e-015 -2.9096e-015 -3.1002e-015 -3.9936e-015 -4.7010e-015 -1.5600e-015
-3.4771e-015 -3.6209e-015 -4.1134e-015 -2.9076e-015 -1.2543e-015 -2.9550e-015 -4.0781e-015 -3.6599e-015 -3.6368e-015 -2.7684e-015
-3.7781e-015 -3.6032e-015 -2.8856e-015 -3.0622e-015 -3.5775e-015 -4.0595e-015 -3.7647e-015 -2.8898e-015 -2.2236e-015 -5.3692e-015
-4.1156e-015 -3.6563e-015 -2.1646e-015 -2.8601e-015 -4.4464e-015 -4.7931e-015 -3.8243e-015 -2.2995e-015 -9.1360e-016 -7.5419e-015
-2.4428e-015 -3.1050e-015 -5.2579e-015 -4.5428e-015 -2.6968e-015 -1.5184e-015 -2.7300e-015 -5.1960e-015 -7.3202e-015  2.6132e-015
```

- Propiedades de los residuos MCO

Los residuos MCO cumplen una serie de propiedades que se enumeran y se comprueban a continuación:

P1. La suma de los residuos es nula

$\sum_{t=1}^{T} e_t = 0 \Rightarrow \iota' e = 0$ donde $\iota' = (1 \quad 1 \quad \ldots \quad 1)_{1xT}$

```
# Suma residual nula
? genr se = sum(E)
Se ha generado el escalar se = -2.63967e-012
```

Consecuencia: La suma de los valores observados del regresando es igual a la suma de sus valores estimados ⇒ La media del regresando coincide con la media de su valor estimado.

$\sum_{t=1}^{T} Y_t = \sum_{t=1}^{T} \hat{Y}_t \Rightarrow \bar{Y} = \bar{\hat{Y}}$

```
# Suma de los valores observados igual a la suma de los valores estimados
? genr sy = sum(Y)
Se ha generado el escalar sy = 103

? genr sye = sum(YE)
Se ha generado el escalar sye = 103

# Media de los valores observados igual a la media de los valores estimados
? genr ym = mean(Y)
Se ha generado el escalar ym = 10.3
```

```
? genr yem = mean(YE)
Se ha generado el escalar yem = 10.3
```

P2. Los residuos y los regresores presentan incorrelación muestral

$$X'e = 0_{(K+1)x1} \implies \sum_{t=1}^{T} X_{it} e_t = 0$$

```
# Incorrelación entre residuos y regresores
? matrix XE = X' * E
Se ha generado la matriz XE

? print XE
XE (3 x 1)
-2.6397e-012
-6.7777e-012
-2.7423e-009
```

P3. No existe correlación muestral entre la variable explicada estimada y los residuos

$$\hat{Y}'e = 0 \implies \sum_{t=1}^{T} \hat{Y}_t e_t = 0$$

```
# Incorrelación entre residuos y valores estimados del regresando
? genr YEE = YE' * E
Se ha generado el escalar YEE = -2.6876e-011
```

P4. El hiperplano de regresión ajustado por MCO pasa por el punto de valores medios $(\bar{Y}, \bar{X}_1, \bar{X}_2, \ldots, \bar{X}_K)$

```
# La recta de regresión pasa por (ym, x1m, x2m x3m x4m)
? genr ymc = b[1] + b[2] * mean(P)+ b[3] * mean(RF)
Se ha generado el escalar ymc = 10.3
```

3.11.2. Interpretación de coeficientes

Los estimadores de los coeficientes de regresión parcial dependen de las unidades de medida de las variables a las que acompañan (por lo que no son directamente comparables). Por ello, que para enriquecer la interpretación de los resultados de la estimación del modelo resulta conveniente calcular estimadores adimensionales.

- Estimadores de los coeficientes estandarizados

$$b_i^* = b_i \frac{s_{X_i}}{s_Y}$$

```
# Coeficientes estandarizados
? genr b0stand = b[1] * sqrt(var(const)) / sqrt(var(C))
Se ha generado el escalar b0stand = 0

? genr b1stand = b[2] * sqrt(var(P)) / sqrt(var(C))
Se ha generado el escalar b1stand = -0.614544

? genr b2stand = b[3] * sqrt(var(RF)) / sqrt(var(C))
Se ha generado el escalar b2stand = 0.661434
```

De los resultados obtenidos se deduce que:

- Manteniendo constante el nivel de renta, ante un incremento de una unidad de desviación estándar en el precio del litro de aceite de oliva, la demanda de aceite disminuye en 0.614544 unidades de desviación estándar.
- Manteniendo constante el precio del litro de aceite de oliva, ante un incremento de una unidad de desviación estándar en el nivel de renta de las familias, la demanda de aceite aumenta en 0.661434 unidades de desviación estándar.

Se puede concluir que la variable que tiene mayor peso en la explicación del consumo mensual de aceite por las familias (C) es el nivel de ingresos mensuales (RF) ya que es la variable con el mayor estimador del coeficiente estandarizado en valor absoluto.

Además, si los regresores son ortogonales, el cuadrado de los coeficientes beta mide la varianza del regresando explicada por esa variable y, por tanto, el coeficiente de determinación vendrá dado por la suma de los cuadrados de dichos coeficientes beta. No obstante, no es habitual trabajar con regresores ortogonales y, por tanto, es imposible aislar del todo el efecto de cada uno de ellos.

```
# Se calculan los cuadrados de los estimadores de los coeficientes beta
? genr b1stand2 = b1stand^2
Se ha generado el escalar b1stand2 = 0.377664

? genr b2stand2 = b2stand^2
Se ha generado el escalar b2stand2 = 0.437495

? genr emulti = 2 * b[2] * b[3] * cov(P,RF) / var(C)
Se ha generado el escalar emulti = 0.1587

? genr R2cemulti = b1stand2 + b2stand2 + emulti
Se ha generado el escalar R2cemulti = 0.973859
```

El R^2 indica que aproximadamente el 97.38% de las variaciones en el consumo mensual de aceite por las familias, son explicadas por las variaciones en el precio del aceite y en el nivel de ingresos mensuales. Además de ese 97% de variaciones en el consumo mensual de aceite por las familias, el 37.77% son explicadas por variaciones en el precio del aceite, el 43.74% son explicadas por variaciones en el nivel de ingresos y el 15.87 % son explicadas por variaciones conjuntas en el nivel de precios y en el nivel de ingresos, dado que ambas variables están correlacionadas:

```
? genr corPRF = corr(P,RF)
Se ha generado el escalar corPRF = -0.195213
```

Otra forma de calcular los estimadores de los coeficientes beta es realizando la estimación del modelo con todas sus variables estandarizadas:

$$\frac{C_t - \bar{C}}{S_C} = \beta_1^* \frac{C_t - \bar{C}}{S_C} + \beta_2^* \frac{RF_t - \overline{RF}}{S_{RF}} + \varepsilon_t^*$$

```
# Variables estandarizadas
? genr Cs = (C - mean(C)) / sqrt(var(C))
Se ha generado la serie Cs (ID 10)

? genr Ps = (P - mean(P)) / sqrt(var(P))
Se ha generado la serie Ps (ID 11)

? genr RFs = (RF - mean(RF)) / sqrt(var(RF))
Se ha generado la serie RFs (ID 12)

# Estimación del modelo con variables estandarizadas
? ols Cs Ps RFs --simple-print

MCO, usando las observaciones 1-10
Variable dependiente: Cs

             Coeficiente   Desv. Típica   Estadístico t   Valor p
  ---------------------------------------------------------------
  Ps          -0.614544     0.0582843        -10.54       5.71e-06 ***
  RFs          0.661434     0.0582843         11.35       3.28e-06 ***

SCR = 0.235267, R-cuadrado = 0.973859
```

El usuario debe tener en cuenta que Gretl denota la SCE (Suma de Cuadrados de los Errores o Residuos) como SCR. En este manual se utiliza SCR para denominar a la Suma de Cuadrados de Regresión.

Cuando las variables de un modelo están centradas, la ordena en el origen se anula y, dado que las variables estandarizadas no son más que las variables centradas divididas por sus desviaciones estándar, no será necesario introducir ordenada en el origen en el modelo:

```
? ols Cs const Ps RFs --simple-print
```

```
MCO, usando las observaciones 1-10
Variable dependiente: Cs

             Coeficiente   Desv. Típica   Estadístico t   Valor p
  ---------------------------------------------------------------
  const       0.000000     0.0579738        0.0000       1.0000
  Ps         -0.614544     0.0623085       -9.863        2.34e-05 ***
  RFs         0.661434     0.0623085       10.62         1.44e-05 ***

SCR = 0.235267, R-cuadrado = 0.973859
```

Se observa que cuando se incluye el regresor ficticio, el estimador de la ordenada en el origen es cero y la probabilidad de rechazar que ese sea su verdadero valor es uno. Otra evidencia de que realmente la ordenada en el origen es cero, es que el coeficiente de determinación del modelo con ordenada en el origen coincide con el coeficiente de determinación bruto del modelo sin ordenada en el origen.

Además, si se comparan los resultados de la estimación del modelo con y sin ordenada en el origen, se obtienen los mismos resultados para los estimadores de los coeficientes beta de las variables explicativas y lo único que varia son los estimadores de las desviaciones estándar, debido a que los grados de libertad en ambos modelos es diferente (en el primero se incluyen dos regresores mientras que en el segundo se incluyen tres). Lo que si resulta relevante y pone de manifiesto que los estimadores de los coeficientes de las variables estandarizadas son directamente comparables es que sus desviaciones típicas estimadas coinciden.

- Estimadores de las elasticidades en media

$$\hat{E}_i = b_i \frac{\bar{X}_i}{\bar{Y}}$$

```
# Elasticidades en media
? genr EP = b[2] * mean(P) / mean(C)
Se ha generado el escalar EP = -0.728034

? genr ERF = b[3] * mean(RF) / mean(C)
Se ha generado el escalar ERF = 0.923513
```

De los resultados obtenidos se deduce que:

- Manteniendo constante el nivel de renta, ante un incremento de un 1% en el precio del litro de aceite de oliva, la demanda de aceite, disminuye un 0.728034%.
- Manteniendo constante el precio del litro de aceite de oliva, ante un incremento de un 1% en el nivel de renta de las familias, la demanda de aceite aumenta en 0.923513%.

Otra forma de calcular los estimadores de las elasticidades es con la estimación del modelo con todas sus variables tomando logaritmos:

$$\ln C_t = \beta_0 + E_1 \ln P_t + E_2 \ln RF_t + \varepsilon_t^*$$

```
# Variables en términos logarítmicos
? genr LC = log(C)
Se ha generado la serie LC (ID 13)

? genr LP = log(P)
Se ha generado la serie LP (ID 14)

? genr LRF = log(RF)
Se ha generado la serie LRF (ID 15)

# Estimación del modelo con variables en términos logaritmicos
? ols LC const LP LRF --simple-print

MCO, usando las observaciones 1-10
Variable dependiente: LC

             Coeficiente   Desv. Típica   Estadístico t   Valor p
```

```
  ---------------------------------------------------------------
```

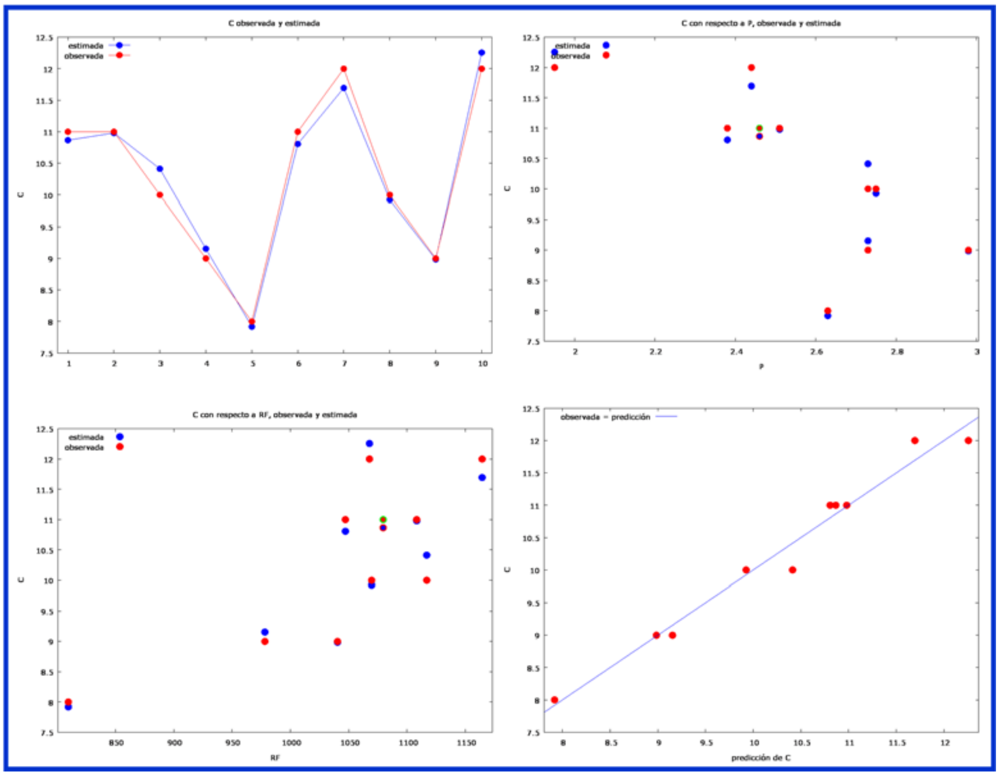

Ilustración 3-1. Gráficos de valores observados y estimados del regresando.

```
  const        -3.47164      0.701716        -4.947      0.0017    ***
  LP           -0.665544     0.0854378       -7.790      0.0001    ***
  LRF           0.923194     0.0981724        9.404      3.20e-05  ***

SCR = 0.00593285, R-cuadrado = 0.963157
```

Las discrepancias existentes con los resultados anteriores son debidas a que en el modelo logarítmico se obtienen estimadores de las elasticidades mientras que con la primera aproximación se obtienen estimadores de las elasticidades medias.

3.11.3. Análisis gráfico

Para hacerse una idea de la “calidad” de la estimación puede resultar interesante la representación gráfica de los valores del regresando y del regresando estimado (véase Ilustración 3-1), así como de los residuos (véase Ilustración 3-2), ya sea respecto al orden de las observaciones o de las variables observables que intervienen en el modelo. La forma más cómoda de acceder a estos gráficos es a través del menú **Gráficos** de la **Ventana Modelo**, en la que se podrán modificar de forma sencilla los atributos de los mismos. Se puede acceder a esta **Ventana Gráfico** ejecutando el correspondiente comando **gnuplot** desde la **Consola Gretl**.

Si el comando **gnuplot** se ejecuta desde un fichero de comandos, no se tiene acceso directamente al gráfico, sino que Gretl en la **salida de guión** informa de la ruta donde se encuentra el fichero que hay que ejecutar para poder acceder a dicho gráfico (por lo que resulta más complicado). En este caso, para modificar sus atributos, el usuario tendrá que estar familiarizado con el **programa gráfico gnuplot**, mientras que si se hace a través de la **Ventana Gráfico**, las modificaciones resultan mucho más sencillas.

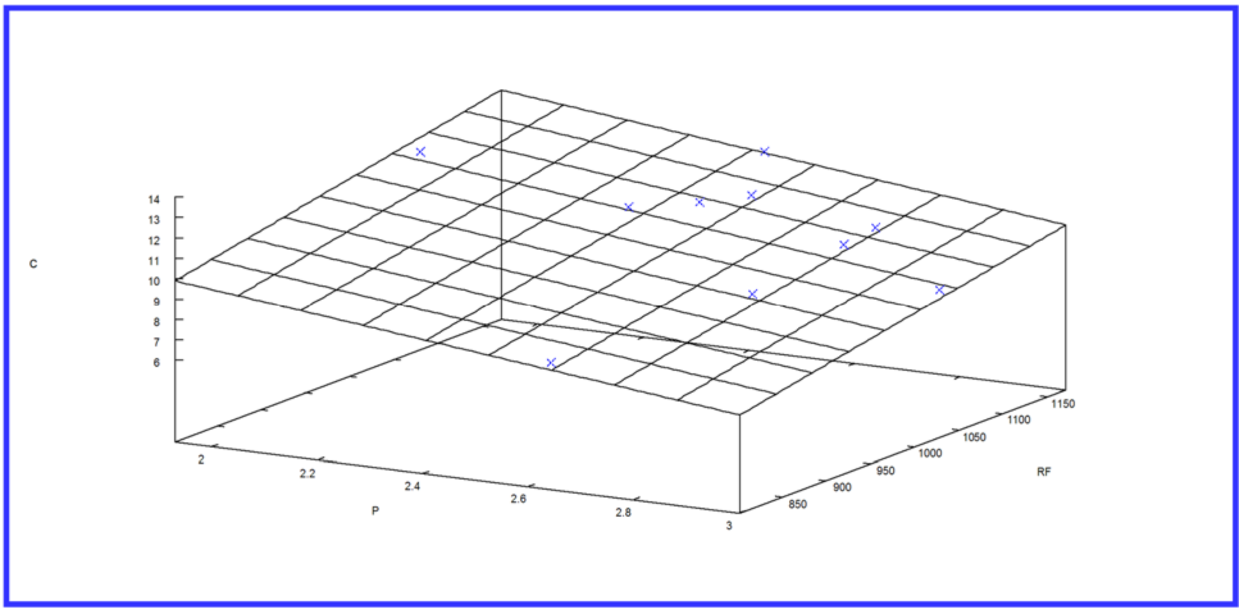

Ilustración 3-3 Gráfico de valores estimados y observados del regresando respecto a las variables explicativas.

Otra forma de construir gráficos es utilizando el comando **graph**. Los gráficos así obtenidos, al ser de texto plano son de peor calidad, pero quizás sean los más apropiados para utilizar en los ficheros de comandos y, recurrir a los de alta calidad, únicamente para los gráficos que deban ir en el informe final del proyecto.

Los gráficos de la Ilustración 3-2 se podrían obtener a partir de la ejecución de los comandos:

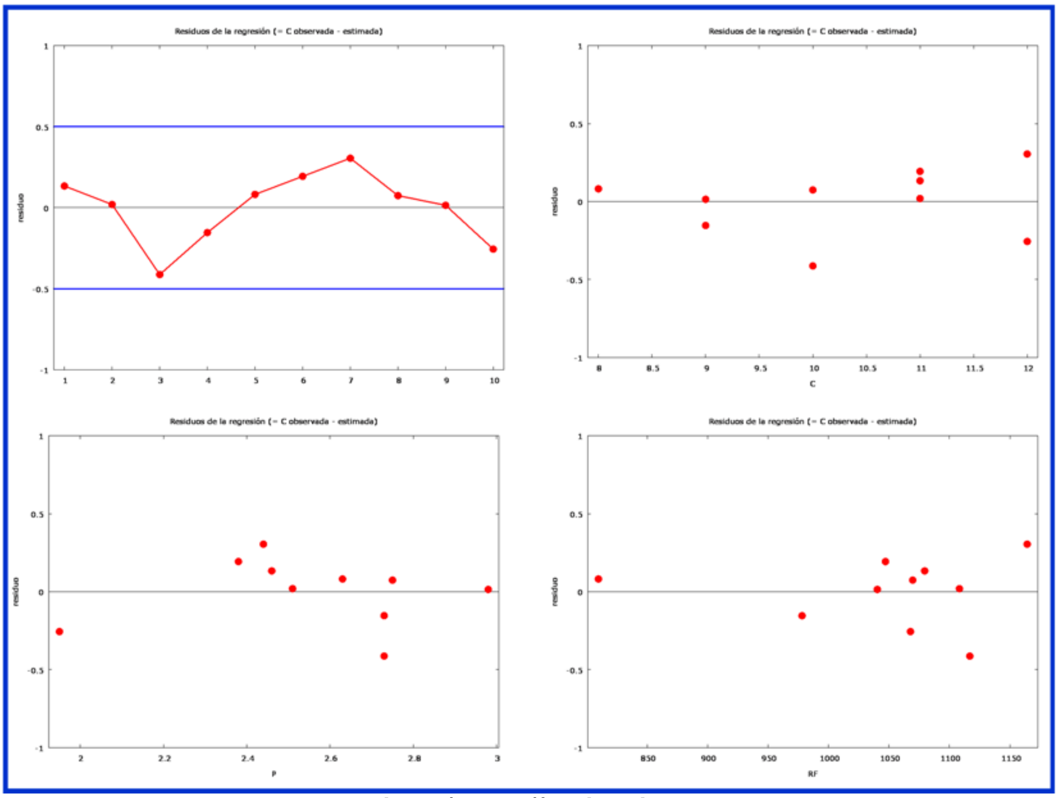

Ilustración 3-2. Gráficos de residuos.

```
# Se genera la serie o
? series o = time
```

```
Se ha generado la serie o (ID 17)

# Gráfico de dispersión de los residuos
? textplot e o

  e
   0.305141  |                                      o
             |
             |
   0.185453  +                               o
             |
             |o
             |                         o                   o
             |
        0.0  +------o--------------------------------------------o-------
             |
             |
             |
             |                   o
  -0.213505  +
             |
             |                                                       o
             |
             |
  -0.412985  +             o
             |+---------+---------+---------+---------+---------+---------+
              1                             o                             10


# Gráfico de dispersión de los residuos
? gnuplot e o
escribió c:\proyecto01\gpttmp01.plt

# Gráfico de dispersión de los residuos respecto a la variable "C"
# textplot e C

# Gráfico de dispersión de los residuos respecto a la variable "P"
# textplot e P

# Gráfico de dispersión de los residuos respecto a la variable "RF"
# textplot e RF

# Gráfico de dispersión de los residuos respecto a la variable "C"
# gnuplot e C

# Gráfico de dispersión de los residuos respecto a la variable "P"
# gnuplot e P

# Gráfico de dispersión de los residuos respecto a la variable "RF"
# gnuplot e RF
```

Cuando se ejecuta un comando **textplot** aparece un gráfico de texto plano y cuando se ejecuta un comando gnuplot aparece la ruta donde se encuentra el fichero correspondiente a dicho gráfico (se ha hecho únicamente para el primer gráfico, pero si se está interesado en obtener los restantes, se debe ejecutar el comando correspondiente y obviar el simbolo # que aparece al principio de la línea del comando).

3.11.4. Estimación MCO del modelo sin ordenada en el origen

```
# Modelo sin ordenada en el origen con anova
? ols C P RF --anova

Modelo 2: MCO, usando las observaciones 1-10
Variable dependiente: C

             Coeficiente   Desv. Típica   Estadístico t   Valor p
  ---------------------------------------------------------------
  P          -1.53921       0.502789        -3.061       0.0156   **
  RF          0.0135517     0.00122794      11.04        4.05e-06 ***

Media de la vble. dep.  10.30000   D.T. de la vble. dep.   1.337494
Suma de cuad. residuos  2.918482   D.T. de la regresión    0.603995
```

```
R-cuadrado            0.997290   R-cuadrado corregido   0.996951
F(2, 8)               1472.110   Valor p (de F)         5.39e-11
Log-verosimilitud    -8.031779   Criterio de Akaike     20.06356
Criterio de Schwarz   20.66873   Crit. de Hannan-Quinn  19.39969
```

Obsérvese que aún cuando el comando **ols** se ha ejecutado con la opción **--anova**, dado que el modelo carece de ordenada en el origen, Gretl no muestra dicha tabla.

- Interpretación de resultados

El modelo estimado por MCO es:

$$\hat{C}_t = -\underset{(0.502789)}{1.53921} P_t + \underset{(0.00122794)}{0.0135517} RF_t$$

$$t_1 = -3.061 \quad t_2 = 11.04$$

$$R^2 = 0.99729$$

De los resultados obtenidos se deduce que:

- Manteniendo constante el nivel de renta, ante un incremento de un euro en el precio del litro de aceite de oliva, la demanda de aceite (medida en términos de consumo mensual por familia), disminuye en 1.53921 litros.
- Manteniendo constante el precio del litro de aceite de oliva, ante un incremento de un euro en el nivel de renta de las familias, la demanda de aceite aumenta en 0.0135517 litros.

Dado que el modelo está formulado sin ordenada en el origen, bajo la denominación de coeficiente de determinación se calcula el coeficiente de determinación bruto, que aunque está acotado entre 0 y 1, carece de la interpretación habitual.

3.11.4.1. Resultados de la estimación utilizando álgebra matricial

En este epígrafe, con la finalidad de entender las diferencias de cálculo entre el modelo con y sin ordenada en el origen, se calculan de forma detallada los resultados de la estimación del modelo.

- Vector de estimadores

Para obtener los estimadores mínimo cuadráticos ordinarios del vector de parámetros es necesario conocer los elementos de la matriz X'X y del vector X'Y, que se obtendrán a partir de los datos de la muestra:

$$b=(X'X)^{-1}X'Y=\begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_K \end{pmatrix}_{Kx1}$$

donde:

- X'X es la matriz de orden KxK, que contiene los sumatorios de los productos cruzados de las variables explicativas:

$$X'X=\begin{pmatrix} \sum_{t=1}^{T} X_{1t}^2 & \sum_{t=1}^{T} X_{1t}X_{2t} & \cdots & \sum_{t=1}^{T} X_{1t}X_{Kt} \\ \sum_{t=1}^{T} X_{2t}X_{1t} & \sum_{t=1}^{T} X_{2t}^2 & \cdots & \sum_{t=1}^{T} X_{2t}X_{Kt} \\ \cdots & \cdots & \ddots & \cdots \\ \sum_{t=1}^{T} X_{Kt}X_{1t} & \sum_{t=1}^{T} X_{Kt}X_{2t} & \cdots & \sum_{t=1}^{T} X_{Kt}^2 \end{pmatrix}_{KxK}$$

- X'Y es el vector de orden Kx1, que contiene los sumatorios de los productos cruzados del regresando con las variables explicativas:

$$X'Y=\begin{pmatrix} \sum_{t=1}^{T} Y_t X_{1t} \\ \sum_{t=1}^{T} Y_t X_{2t} \\ \cdots \\ \sum_{t=1}^{T} Y_t X_{Kt} \end{pmatrix}_{Kx1}$$

```
# Se genera la matriz de diseño (Xso)
? matrix Xso = {P, RF}
Se ha generado la matriz Xso

# Se muestra la matriz Xso que, en este caso, es de orden 10x2
? print Xso
Xso (10 x 2) [t1 = 1, t2 = 10]
      2.4600       1079.5
      2.5100       1108.2
      2.7300       1116.8
      2.7300       978.09
      2.6300       809.62
      2.3800       1047.1
      2.4400       1164.3
      2.7500       1069.6
      2.9800       1040.3
      1.9500       1067.8
```

```
# Se calcula y se muestra la matriz de productos cruzados de las variables explicativas
? matrix  XTXso = Xso'* Xso
Se ha generado la matriz XTXso

? print XTXso
XTXso (2 x 2)
      66.038       26743.
      26743.   1.1072e+007

# Se calcula y se muestra la inversa de X'Xso
? genr XTXsoI = inv(XTXso)
Se ha generado la matriz XTXsoI

? print XTXsoI
XTXsoI (2 x 2)
     0.69295   -0.0016738
  -0.0016738   4.1332e-006
```

```
# Se calcula y se muestra el vector de productos cruzados de las variables explicativas y el regresando
? matrix XTsoY = Xso'* Y
Se ha generado la matriz XTsoY

? print XTsoY
XTsoY (2 x 1)
      260.76
 1.0888e+005

# Se calcula bso
? matrix bso = XTXsoI * XTsoY
Se ha generado la matriz bso

? print bso
bso (2 x 1)
     -1.5392
    0.013552
```

- Regresando estimado

$\hat{Y} = \mathrm{Xb}$

```
# Se calcula el regresando estimado
? matrix YEso = Xso * bso
Se ha generado la matriz YEso
? print YEso
YEso (10 x 1) [t1 = 1, t2 = 10]

      10.843
      11.155
      10.933
      9.0527
      6.9236
      10.526
      12.023
      10.262
      9.5117
      11.468
```

Si lo que interesa obtener es la estimación para una única observación del regresando, se puede utilizar:

$\hat{Y}_t = X_t'\mathrm{b} = \mathrm{b}_1 X_{1t} + \cdots + \mathrm{b}_K X_{Kt}$

```
# Se calcula y se muestra el valor estimado para la primera observación
? genr YEso1 = bso[1] * Xso[1,1] + bso[2] * Xso[1,2]
Se ha generado el escalar YEso1 = 10.8427
```

Para generar la serie que contiene los valores estimados del regresando se utiliza :

```
# Se genera y se muestra la serie que recoge las observaciones del regresando estimado (ye)
? series yeso = Xso * bso
Se ha generado la serie yeso (ID 17)

? print yeso
Muestra actual:  1 - 10 (n = 10)
10.8427  11.1548  10.9330  9.05273  6.92359  10.5264  12.0231  10.2621
9.51165  11.4685
```

- Residuos

$e = Y - \hat{Y}$

```
# Se calcula el vector de residuos
? matrix Eso = Y - YEso
Se ha generado la matriz Eso
? print Eso
Eso (10 x 1) [t1 = 1, t2 = 10]

     0.15727
    -0.15484
    -0.93302
   -0.052727
      1.0764
```

```
     0.47362
   -0.023102
    -0.26206
    -0.51165
     0.53151
```

Para obtener un residuo para una observación concreta:

$$e_t = Y_t - \hat{Y}_t$$

```
# Se calcula y se muestra el valor estimado para la primera observación
? genr eso1 = Y[1] - YEso[1]
Se ha generado el escalar eso1 = 0.157273
```

Para obtener la serie de residuos:

```
# Se genera y se muestra la serie de residuos
? series eso = Y - YEso
Se ha generado la serie eso (ID 18)
? print eso
Muestra actual:  1 - 10 (n = 10)
 0.157273  -0.154836  -0.933024  -0.0527265   1.07641   0.473617  -0.0231017
-0.262058  -0.511652   0.531507
```

- Suma de Cuadrados de Errores

$$SCE = e'e = \sum_{t=1}^{T} e_t^2 = Y'Y - b'X'Y$$

```
# Se calcula la SCE
? genr SCEso = Eso' * Eso
Se ha generado el escalar SCEso = 2.91848
```

- Estimador de la varianza de la perturbación

$$S^2 = \frac{SCE}{T-K}$$

```
# Se calcula el estimador de la varianza de la perturbación
? genr s2so = SCEso/ (T - k)
Se ha generado el escalar s2so = 0.36481

# Se calcula el estimador de la desviación típica de la perturbación
? genr sso = sqrt(s2so)
Se ha generado el escalar sso = 0.603995
```

- Matriz de varianzas-covarianzas estimada de los EMCO

$$\hat{V}(b) = S^2(X'X)^{-1} = \begin{pmatrix} S_{b_1}^2 & S_{b_1b_2} & \cdots & S_{b_1b_K} \\ S_{b_2b_1} & S_{b_2}^2 & \cdots & S_{b_2b_K} \\ \cdots & \cdots & \ddots & \cdots \\ S_{b_Kb_1} & S_{b_Kb_2} & \cdots & S_{b_K}^2 \end{pmatrix}_{KxK}$$

```
# Se calcula y se muestra la matriz de varianzas-covarianzas estimada de los estimadores
? matrix Vbso = s2so * inv(Xso' * Xso)
Se ha generado la matriz Vbso

? print Vbso
Vbso (2 x 2)
     0.25280  -0.00061061
 -0.00061061  1.5078e-006

# Se calcula y se muestra el vector que contiene las varianzas estimadas de los EMCO
? matrix s2bso = diag(Vbso)
Se ha generado la matriz s2bso

? print s2bso
s2bso (2 x 1)
     0.25280
 1.5078e-006
```

```
# Se calcula y se muestra el vector que contiene las desviaciones típicas estimadas de los EMCO
? matrix sbso = sqrt(s2bso)
Se ha generado la matriz sbso

? print sbso
sbso (2 x 1)
     0.50279
   0.0012279
```

- Ratios t de los EMCO

$$t_i = \frac{b_i}{S_{b_i}}$$

```
# Se calcula las ratios t
? genr t1so = bso[1] / sbso[1]
Se ha generado el escalar t1so = -3.06135

? genr t2so = bso[2] / sbso[2]
Se ha generado el escalar t2so = 11.0361
```

- Tabla ANOVA respecto a la media y Tabla ANOVA respecto al origen

Aunque Gretl para modelos sin ordenada en el origen no proporciona ninguna de las tablas de Análisis de la Varianza, se ha considerado importante construirlas para que se pueda entender como se calcula la bondad de ajuste para estos modelos y para que quede claro que aunque el coeficiente calculado está acotado, no tiene la interpretación habitual, puesto que su cálculo se basa en la Tabla ANOVA respecto al origen (véase **Tabla 3-4**) y no en la Tabla ANOVA respecto a la media (véase **Tabla 3-3**).

```
# ANOVA respecto a la media
# Se calcula la SCT
? genr SCT = sum((y - mean(y))^2)
Se ha reemplazado el escalar SCT = 16.1

# Se calcula la SCR
? genr SCRso = sum((yeso - mean(yeso))^2)
Se ha generado el escalar SCRso = 19.3814

# Se calcula la SCE
? genr SCEso = Eso' * Eso
Se ha reemplazado el escalar SCEso = 2.91848
```

Tabla 3-3. Tabla **ANOVA** respecto a la media.

Análisis de Varianza:

	Sumas de cuadrados	gl	Media de cuadrados
Regresión $SCR = \sum_{t=1}^{T}(\hat{Y}_t - \bar{Y})^2$	19.3814	2	…
Residuo $SCE = \sum_{t=1}^{T} e_t^2$	2.91848	8	…
Total $SCT = \sum_{t=1}^{T}(Y_t - \bar{Y})^2$	16.1	9	…

R^2 = 19.3814 / 16.1 = 1.2038

…

```
# ANOVA respecto al origen
? genr scy = sum(y^2)
Se ha reemplazado el escalar scy = 1077

? genr scyeso = sum(yeso^2)
```

```
Se ha generado el escalar scyeso = 1074.08

? genr sceso = sum(eso^2)
Se ha generado el escalar sceso = 2.91848
```

Tabla 3-4. Tabla **ANOVA** respecto al origen.

Análisis de Varianza:

		Sumas de cuadrados	gl	Media de cuadrados
Regresión	$\sum_{t=1}^{T}(\hat{Y}_t-0)^2=\sum_{t=1}^{T}\hat{Y}_t^2$	1074.08	2	537.04
Residuo	$\sum_{t=1}^{T}(e_t-0)^2=\sum_{t=1}^{T}e_t^2=SCE$	2.91848	8	0.36481
Total	$\sum_{t=1}^{T}(Y_t-0)^2=\sum_{t=1}^{T}Y_t^2$	1077	10	107.7

R^2 = 1074.08 / 1077 = 0.9972

F(2, 8) = 537.04 / 0.36481 = 1472.110 [Valor p 2.74605e-012]

- Cálculo del coeficiente de determinación

Como el modelo está formulado sin ordenada en el origen:

$$R^2=\frac{SCR}{SCT}\neq 1-\frac{SCE}{SCT}$$

```
# Cálculo del coeficiente de determinación
? genr R2soc1 = SCRso/SCT
Se ha generado el escalar R2soc1 = 1.20381

? genr R2soc2 = 1 - SCEso/SCT
Se ha generado el escalar R2soc2 = 0.818728
```

Se debe tener en cuenta que si en un modelo sin ordenada en el origen se calcula el coeficiente de determinación como $SCR\ /\ SCT$, aunque la cota inferior seguirá siendo cero, dejará de estar acotado superiormente (como en este caso, podra ser superior a uno).

Si se calcula dicho coeficiente como $1-(SCE\ /\ SCT)$, seguirá teniendo cota superior uno, pero ya no estará acotado inferiormente (pudiendo en este caso, tomar valores negativos).

- Coeficiente de determinación bruto

Se puede comprobar que el valor que aparece bajo la denominación **R-cuadrado** no es el coeficiente de determinación propiamente dicho sino el coeficiente de determinación bruto: **`R-cuadrado 0.997290`**, cálculo que se basa en la Tabla ANOVA respecto al origen

```
# Cálculo del coeficiente de determinación bruto
? genr R2RAWso = scyeso/scy
Se ha generado el escalar R2RAWso = 0.99729

? genr R2RAWso = 1 - sceso/scy
Se ha reemplazado el escalar R2RAWso = 0.99729
```

- Coeficiente de determinación corregido o ajustado

$$\bar{R}^2=1-\frac{\frac{SCE}{T-K}}{\frac{\sum_{t=1}^{T}Y_t^2}{T-1}}$$

```
# Se calcula el coeficiente de determinación ajustado
? genr R2Aso = 1 - ((SCEso / (T - k)) / (scy / (T - 1)))
```

```
Se ha generado el escalar R2Aso = 0.996951
```

- Propiedades de los residuos MCO

Cuando el modelo está formulado sin ordenada en el origen, los residuos MCO cumplen las siguientes propiedades:

P1. La suma de los residuos no es nula

$\sum_{t=1}^{T} e_t \neq 0 \implies \iota' e \neq 0$ donde $\iota' = (1 \quad 1 \quad \ldots \quad 1)_{1xT}$

```
# Suma residual no es nula
? genr se = sum(eso)
Se ha reemplazado el escalar se = 0.301405
```

Consecuencia: La suma de los valores observados del regresando no es igual a la suma de sus valores estimados ⇒ La media del regresando no coincide con la media de su valor estimado.

$\sum_{t=1}^{T} Y_t \neq \sum_{t=1}^{T} \hat{Y}_t \implies \bar{Y} \neq \bar{\hat{Y}}$

```
# Suma de los valores observados no es igual a la suma de los valores estimados
? genr sy = sum(y)
Se ha reemplazado el escalar sy = 103

? genr syeso = sum(yeso)
Se ha generado el escalar syeso = 102.699

# Media de los valores observados no es igual a la media de los valores estimados
? genr ym = mean(y)
Se ha reemplazado el escalar ym = 10.3

? genr yesom = mean(yeso)
Se ha generado el escalar yesom = 10.2699
```

P2. Los residuos y las variables explicativas presentan incorrelación muestral

$X'e = 0_{Kx1} \implies \sum_{t=1}^{T} X_{it} e_t = 0$

```
# Incorrelación entre residuos y variables explicativas
? matrix XsoEso = Xso' * Eso
Se ha generado la matriz XsoEso

? print XsoEso
XsoEso (2 x 1)
 7.3632e-013
 3.0567e-010
```

P3. No existe correlación muestral entre la variable explicada estimada y los residuos

$\hat{Y}'e = 0 \implies \sum_{t=1}^{T} \hat{Y}_t e_t = 0$

```
# Incorrelación entre residuos y valores estimados
? genr YEsoEso = YEso' * Eso
Se ha generado el escalar YEsoEso = 3.00679e-012
```

P4. El hiperplano de regresión ajustado por MCO no pasa por el punto de valores medios

El hiperplano de regresión ajustado no pasa por el punto $(\bar{Y}, \bar{X}_1, \bar{X}_2, \ldots, \bar{X}_K)$ sino que pasa por el punto $(\bar{\hat{Y}}, \bar{X}_1, \bar{X}_2, \ldots, \bar{X}_K)$:

```
# La recta de regresión pasa por (yem, x1m, x2m x3m x4m)
? genr yesomc = bso[1] * mean(P)+ bso[2] * mean(RF)
Se ha generado el escalar yesomc = 10.2699

? print yesomc yesom ym

         yesomc =  10.269859
          yesom =  10.269859
             ym =  10.300000
```

3.11.5. Formas funcionales alternativas

Debe tenerse en cuenta, que la interpretación de los coeficientes estimados es distinta dependiendo de la forma funcional del modelo especificado y que el coeficiente de determinación, tan sólo, se puede utilizar para comparar modelos que tengan la misma variable dependiente, el mismo número de regresores e identica forma funcional. En este caso, no se podría utilizar como criterio de selección de modelos, puesto que no serían directamente comparables, por ello, para poder realizar dicha comparación se ha calculado el Coeficiente de determinación equivalente:

$$R^2_{equi} = 1 - \frac{SCE_{en\ unidades\ de\ Y}}{SCT_{en\ unidades\ de\ Y}}$$

```
# Transformación de las variables
? genr LC = log(C)
Se ha generado la serie LC (ID 7)

? genr LP = log(P)
Se ha generado la serie LP (ID 8)

? genr LRF = log(RF)
Se ha generado la serie LRF (ID 9)

? genr PI = 1 / P
Se ha generado la serie PI (ID 10)

? genr RFI = 1 / RF
Se ha generado la serie RFI (ID 11)
```

- ***Modelo lineal o Modelo Lin-Lin***

$$C_t = \beta_0 + \beta_1 P_t + \beta_2 RF_t + \varepsilon_t$$

```
# Modelo 1: Modelo Lin-Lin
? ols C const P RF --simple-print

MCO, usando las observaciones 1-10
Variable dependiente: C

             Coeficiente   Desv. Típica   Estadístico t   Valor p
  ---------------------------------------------------------------
  const        8.28657      1.28569          6.445       0.0004   ***
  P           -2.93379      0.297456        -9.863       2.34e-05 ***
  RF           0.00907529   0.000854912     10.62        1.44e-05 ***

SCR = 0.420867, R-cuadrado = 0.973859
```

$b_1 = \partial C/\partial P \Longrightarrow b_1 = \Delta C/\Delta P$ → Manteniendo constante el nivel de renta de las familias, si se incrementa en 1 euro el precio del litro de aceite, la cantidad demandada disminuye en 2.934 litros.

$b_2 = \partial C/\partial RF \Longrightarrow b_2 = \Delta C/\Delta RF$ → Manteniendo constante el precio del litro de aceite, si se incrementa en 1 euro la renta de las familias, la cantidad demandada aumenta en 0.009075 litros.

Dado que la variable dependiente es lineal, el coeficiente de determinación equivalente coincide con el coeficiente de determinación.

```
# Cálculo del coeficiente de determinación equivalente
? series CEM1 = $yhat
Se ha generado la serie CEM1 (ID 12)

? series EEQUIM1 = C - CEM1
Se ha generado la serie EEQUIM1 (ID 13)

? genr SCEEQUIM1 = sum(EEQUIM1^2)
Se ha generado el escalar SCEEQUIM1 = 0.420867

? genr R2EQUIM1 = 1 - SCEEQUIM1 / (sum((C - mean(C))^2))
Se ha generado el escalar R2EQUIM1 = 0.973859
```

- ***Modelo Lin-Log***

$$C_t = \beta_0 + \beta_1 \ln P_t + \beta_2 \ln RF_t + \varepsilon_t$$

```
# Modelo 2: Modelo Lin-Log
? ols C const LP LRF --simple-print

MCO, usando las observaciones 1-10
Variable dependiente: C

             Coeficiente   Desv. Típica   Estadístico t   Valor p
  ----------------------------------------------------------------
  const        -44.0633       7.49873         -5.876       0.0006   ***
  LP            -7.05398      0.913011        -7.726       0.0001   ***
  LRF            8.76808      1.04910          8.358       6.89e-05 ***

SCR = 0.677511, R-cuadrado = 0.957919
```

$b_1 = \partial C / \partial \ln P \Longrightarrow b_1 = \Delta C / (\Delta P / P)$ → Manteniendo constante el nivel de renta de las familias, si se incrementa en un 1% el precio del litro de aceite, la cantidad demandada disminuye en 0.07054 litros.

$b_2 = \partial C / \partial \ln RF \Longrightarrow b_2 = \Delta C / (\Delta RF / RF)$ → Manteniendo constante el precio del litro de aceite, si se incrementa en un 1% la renta de las familias, la cantidad demandada aumenta en 0.08768 litros.

Dado que la variable dependiente es lineal, el coeficiente de determinación equivalente coincide con el coeficiente de determinación.

```
# Cálculo del coeficiente de determinación equivalente
? series CEM2 = $yhat
Se ha generado la serie CEM2 (ID 14)

? series EEQUIM2 = C - CEM2
Se ha generado la serie EEQUIM2 (ID 15)

? genr SCEEQUIM2 = sum(EEQUIM2^2)
Se ha generado el escalar SCEEQUIM2 = 0.677511

? genr R2EQUIM2 = 1 - SCEEQUIM2 / (sum((C - mean(C))^2))
Se ha generado el escalar R2EQUIM2 = 0.957919
```

- ***Modelo Log-Log***

$$\ln C_t = \beta_0 + \beta_1 \ln P_t + \beta_2 \ln RF_t + \varepsilon_t$$

```
# Modelo 3: Modelo Log-Log
? ols LC const LP LRF --simple-print

MCO, usando las observaciones 1-10
Variable dependiente: LC

             Coeficiente   Desv. Típica   Estadístico t   Valor p
  ----------------------------------------------------------------
  const        -3.47164      0.701716         -4.947       0.0017   ***
  LP           -0.665544     0.0854378        -7.790       0.0001   ***
  LRF           0.923194     0.0981724         9.404       3.20e-05 ***

SCR = 0.00593285, R-cuadrado = 0.963157
```

$b_1 = \partial \ln C / \partial \ln P \Longrightarrow b_1 = (\Delta C / C)/(\Delta P / P)$ → Manteniendo constante el nivel de renta de las familias, si se incrementa en un 1% el precio del litro de aceite, la cantidad demandada disminuye en un 0.6655%.

$b_2 = \partial \ln C / \partial \ln RF \Longrightarrow b_2 = (\Delta C / C)/(\Delta \mathrm{RF} / RF)$ → Manteniendo constante el precio del litro de aceite, si se incrementa en un 1% la renta de las familias, la cantidad demandada aumenta en un 0.9232%.

En este caso, los coeficientes estimados miden la elasticidad-precio y la elasticidad-renta de la demanda de aceite de oliva. Como ambas elasticidades son menores que 1 en términos absolutos, se puede concluir que la demanda de aceite de oliva es inelástica tanto respecto al precio como respecto al nivel de renta.

Dado que la variable dependiente no es lineal, el coeficiente de determinación equivalente no coincide con el coeficiente de determinación.

```
# Cálculo del coeficiente de determinación equivalente
? series LCEM3 = $yhat
Se ha generado la serie LCEM3 (ID 16)

? series CEM3 = exp(LCEM3)
Se ha generado la serie CEM3 (ID 17)

? series EEQUIM3 = C - CEM3
Se ha generado la serie EEQUIM3 (ID 18)

? genr SCEEQUIM3 = sum(EEQUIM3^2)
Se ha generado el escalar SCEEQUIM3 = 0.711406

? genr R2EQUIM3 = 1 - SCEEQUIM3 / (sum((C - mean(C))^2))
Se ha generado el escalar R2EQUIM3 = 0.955813
```

- ***Modelo Log-Lin***

$$\ln C_t = \beta_0 + \beta_1 P_t + \beta_2 RF_t + \varepsilon_t$$

```
# Modelo4 log-lin
? ols LC const P RF --simple-print

MCO, usando las observaciones 1-10
Variable dependiente: LC

             Coeficiente    Desv. Típica   Estadístico t   Valor p
  ----------------------------------------------------------------
  const       2.03348       0.121789         16.70         6.76e-07 ***
  P          -0.277068      0.0281771        -9.833        2.39e-05 ***
  RF          0.000953075   8.09834e-05      11.77         7.24e-06 ***

SCR = 0.00377653, R-cuadrado = 0.976548
```

$b_1 = \partial \ln C / \partial P \Longrightarrow b_1 = (\Delta C / C) / \Delta P$ → Manteniendo constante el nivel de renta de las familias, si se incrementa en 1 euro el precio del litro de aceite, la cantidad demandada disminuye en un 27.71%.

$b_2 = \partial \ln C / \partial RF \Longrightarrow b_2 = (\Delta C / C) / \Delta \mathrm{RF}$ → Manteniendo constante el precio del litro de aceite, si se incrementa en 1 euro la renta de las familias, la cantidad demandada aumenta en un 0.09531%.

Dado que la variable dependiente no es lineal, el coeficiente de determinación equivalente no coincide con el coeficiente de determinación.

```
# Cálculo del coeficiente de determinación equivalente
? series LCEM4 = $yhat
Se ha generado la serie LCEM4 (ID 19)

? series CEM4 = exp(LCEM4)
Se ha generado la serie CEM4 (ID 20)

? series EEQUIM4 = C - CEM4
Se ha generado la serie EEQUIM4 (ID 21)

? genr SCEEQUIM4 = sum(EEQUIM4^2)
Se ha generado el escalar SCEEQUIM4 = 0.44362

? genr R2EQUIM4 = 1 - SCEEQUIM4 / (sum((C - mean(C))^2))
Se ha generado el escalar R2EQUIM4 = 0.972446
```

Cuando en una sesión de trabajo se estiman varios modelos con el menú **Modelo** de la **Ventana Principal** y se guardan los resultados que Gretl muestra en la **Ventana Modelo** como icono, se permite crear lo que Gretl denomina **Tabla de modelos**, donde se muestran los resultados de los principales estadísticos obtenidos en las estimaciones realizadas. Para crear la **Tabla de modelos** se debe acceder a la **Ventana vista de iconos** y situarse con el botón derecho del ratón encima del icono correspondiente al modelo cuyos resultados se deseen añadir a dicha tabla y en el menú emergente, seleccionar **Añadir a la tabla de modelos**.

```
Estimaciones de MCO
Variable dependiente: C

              (Modelo 1)   (Modelo 2)

const            8.287**     -44.06**
                (1.286)      (7.499)

P               -2.934**
               (0.2975)

RF           0.009075**
            (0.0008549)

LP                           -7.054**
                            (0.9130)

LRF                           8.768**
                             (1.049)


        n          10           10
R-cuadrado       0.9739       0.9579
       lnL       1.651       -0.7298

Desviaciones típicas entre paréntesis
* indica significativo al nivel del 10 por ciento
** indica significativo al nivel del 5 por ciento
```

En este caso, dado que la variable dependiente no es la misma en todos los modelos, no se van a poder incluir todos los resultados en una única tabla. Se han tenido que utilizar dos tablas: en la primera se recogen los dos modelos cuya variable dependiente es el consumo y en la segunda los dos modelos restantes, que tienen como variable dependiente el logaritmo neperiano del consumo.

```
Estimaciones de MCO
Variable dependiente: LC

              (Modelo 3)    Modelo 4)

const           -3.472**      2.033**

               (0.7017)     (0.1218)

LP             -0.6655**
              (0.08544)

LRF             0.9232**
              (0.09817)

P                            -0.2771**
                            (0.02818)

RF                         0.0009531**
                          (8.098e-05)

PI


RFI


        n          10           10
R-cuadrado       0.9632       0.9765
       lnL       22.96        25.22

Desviaciones típicas entre paréntesis
* indica significativo al nivel del 10 por ciento
```

```
** indica significativo al nivel del 5 por ciento
```

- ¿Qué forma funcional es la más adecuada?

Tabla 3-5. Comparación bondad de ajuste.

	R^2	R^2_{equi}
Modelo Lineal	0.9739	0.9739
Modelo Lin-Log	0.9579	0.9579
Modelo Log-Log	0.9632	0.9558
Modelo Log-Lin	0..9765	0.9725

Para comparar modelos doblemente logarítmicos y semilogarítmicos en la variable dependiente con modelos lineales, no se debe utilizar el coeficiente de determinación, sino el coeficiente de determinación equivalente (véase **Tabla 3-5**).

Al fijarse en el coeficiente de determinación equivalente, se selecciona el modelo lineal, por ser el que presenta un mayor valor para este coeficiente.

Los gráficos que aparecen en la Ilustración 3-4 también pueden ayudar a vislumbrar cual es la forma funcional más adecuada, aún cuando, hay que señalar, que en este ejemplo las diferencias son mínimas.

3.11.6. Predicción

Para realizar una predicción es necesario que en un mismo fichero se encuentren los datos muestrales (estimación) y los datos extramuestrales (predicción). El fichero aceite.gdt contiene 60 observaciones, se han utilizado las 10 primeras para la estimación y se utilizarán las observaciones 11 y 12 para efectuar predicciones.

3.11.6.1. Predicción a través del comando fcast

El comando **fcast** debe ir precedido necesariamente de un comando **ols**, de lo contrario Gretl indicará mediante un mensaje que se está cometiendo un error.

```
? fcast 11 12  --static
No se puede hacer esto: ningún modelo ha sido aún estimado
Error de datos
```

```
? smpl 1 10
```

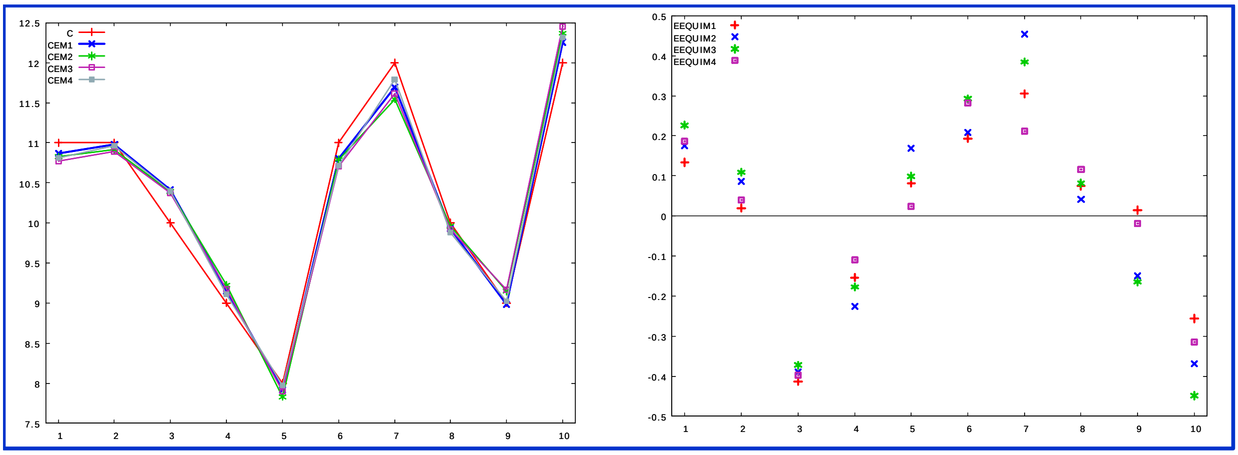

Ilustración 3-4. Gráfico valores estimados del consumo y Gráfico residuos equivalentes.

```
Rango de datos completo: 1 - 60 (n = 60)
Muestra actual: 1 - 10 (n = 10)

# Estimación del modelo
? ols C const P RF --quiet
```

- Predicción puntual

```
? smpl 1 12
```

```
Rango de datos completo: 1 - 60 (n = 60)
Muestra actual: 1 - 12 (n = 12)

? fcast 11 12  --static --no-stats

 Para intervalos de confianza 95%, t(7, .0.025) = 2.365

Observaciones        C    predicción  Desv. Típica     Intervalo de confianza 95%

      11         13.00        13.00        0.341        12.19 -     13.80
      12         12.00        12.10        0.282        11.43 -     12.77
```

Dado que se ha ejecutado el comando **fcast** con las opciones **--static** y **--no-stats**, se obtienen predicciones estáticas puntuales y se suprime la salida de los estadísticos de análisis de la capacidad predictiva.

- Predicción media

```
? fcast 11 12  --mean-y --static --no-stats

 Para intervalos de confianza 95%, t(7, .0.025) = 2.365

Observaciones        C    predicción  Desv. Típica     Intervalo de confianza 95%

      11         13.00        13.00        0.237        12.44 -     13.56
      12         12.00        12.10        0.140        11.77 -     12.43
```

Para obtener predicciones medias puntuales es necesario que el comando **fcast** se ejecute con la opción **--mean-y**.

Obsérvese que la predicción puntual y la predicción media coinciden, lo que varia es el estimador de la desviación típica y, por tanto, los intervalos de confianza, siendo más amplios para la predicción media que para la predicción puntual.

3.11.6.2. Predicción utilizando álgebra matricial

- Regresando predicho

Bajo las hipótesis del MRLNC y asumiendo estabilidad postmuestral, el predictor óptimo en el sentido de minimizar la varianza del error de predicción, viene dado por:

$$\hat{Y}_\tau = X_\tau' b = b_0 + b_1 X_{1\tau} + b_2 X_{2\tau} + \cdots + b_K X_{K\tau} \ \forall\tau = 1, 2, \ldots, N$$

Para realizar una predicción es necesario conocer los valores de los regresores en el período de predicción.

```
# Se define el vector que recoge las observaciones de los regresores para la observación 11
? matrix X11 = {const[11], P[11], RF[11]}
Se ha generado la matriz X11

? print X11
X11 (1 x 3)
      1.0000       2.4600       1314.5

# Se define el vector que recoge las observaciones de los regresores para la observación 12
? matrix X12 = {const[12], P[12], RF[12]}
Se ha generado la matriz X12

? print X12
X12 (1 x 3)
      1.0000       2.3400       1176.9

# Predicción para la observación 11
? genr Cp11 = X11 * b
Se ha generado el escalar Cp11 = 12.9986

# Predicción para la observación 12
? genr Cp12 =  X12 * b
Se ha generado el escalar Cp12 = 12.1022
```

- Error de predicción

$$e_\tau = Y_\tau - \hat{Y}_\tau \ \ \forall \tau = 1, 2, \ldots, N$$

```
# Error de predicción para la observación 11
? genr ep11 = C[11] - Cp11
Se ha generado el escalar ep11 = 0.00143375

# Error de predicción para la observación 12
? genr ep12 = C[12] - Cp12
Se ha generado el escalar ep12 = -0.102223
```

- Varianza estimada del error de la predicción puntual

$$S_{e_\tau}^2 = S^2[1 + X_\tau'(X'X)^{-1}X_\tau] \ \ \forall \tau = 1, 2, \ldots, N$$

```
# Varianza estimada del error de predicción puntual para la observación 11
? genr s2epp11 = s2 * (1 + X11 * inv(X'X) * X11')
Se ha generado el escalar s2epp11 = 0.116251

# Varianza estimada del error de predicción puntual para la observación 12
? genr s2epp12 = s2 * (1 + X12 * inv(X'X) * X12')
Se ha generado el escalar s2epp12 = 0.0796202
```

Al interpretar el resultado, si la varianza del error de predicción es pequeña se puede esperar que los errores de predicción también sean pequeños.

- Desviación típica estimada del error de la predicción puntual

$$S_{e_\tau} = \sqrt{S_{e_\tau}^2} = \sqrt{S^2[1 + X_\tau'(X'X)^{-1}X_\tau]} \ \ \forall \tau = 1, 2, \ldots, N$$

```
# Desviación típica estimada del error de predicción puntual para la observación 11
? genr sepp11 = sqrt(s2epp11)
Se ha generado el escalar sepp11 = 0.340956

# Desviación típica estimada del error de predicción puntual para la observación 12
? genr sepp12 = sqrt(s2epp12)
Se ha generado el escalar sepp12 = 0.28217
```

- Varianza estimada del error de la predicción media

$$S_{e_\tau}^2 = S^2[X_\tau'(X'X)^{-1}X_\tau] \ \ \forall \tau = 1, 2, \ldots, N$$

```
# Varianza estimada del error de predicción media para la observación 11
? genr s2epm11 = s2 * (X11 * inv(X'X) * X11')
Se ha generado el escalar s2epm11 = 0.0561274

# Varianza estimada del error de predicción media para la observación 12
? genr s2epm12 = s2 * (X12 * inv(X'X) * X12')
Se ha generado el escalar s2epm12 = 0.0194963
```

- Desviación típica estimada del error de la predicción media

$$S_{e_\tau} = \sqrt{S_{e_\tau}^2} = \sqrt{S^2[X_\tau'(X'X)^{-1}X_\tau]} \ \ \forall \tau = 1, 2, \ldots, N$$

```
# Desviación típica estimada del error de predicción media para la observación 11
? genr sepm11 = sqrt(s2epm11)
Se ha generado el escalar sepm11 = 0.236912

# Desviación típica estimada del error de predicción media para la observación 12
? genr sepm12 = sqrt(s2epm12)
Se ha generado el escalar sepm12 = 0.139629
```

- Intervalo de confianza para la predicción puntual

$$\hat{Y}_\tau \pm t_{\alpha/2}^{T-K-1} S_{e_\tau} \ \ \forall \tau = 1, 2, \ldots, N$$

```
# Valor critico para una distribución t de Student de 7 grados de libertad (T-k-1) y un nivel de
significación del 5%
? genr talpham = critical(t,7,0.025)
```

```
Se ha generado el escalar talpham = 2.36462

# Intervalo de confianza para un nivel de significación del 5% para la observación 11
? genr LIIPP11 = Cp11 - talpham * sepp11
Se ha generado el escalar LIIPP11 = 12.1923

? genr LSIPP11 = Cp11 + talpham * sepp11
Se ha generado el escalar LSIPP11 = 13.8048

# Intervalo de confianza para un nivel de significación del 5% para la observación 12
? genr LIIPP12 = Cp12 - talpham * sepp12
Se ha generado el escalar LIIPP12 = 11.435

? genr LSIPP12 = Cp12 + talpham * sepp12
Se ha generado el escalar LSIPP12 = 12.7694
```

- Intervalo de confianza para la predicción media

Dado que el predictor del valor esperado coincide con el predictor puntual:

$$\hat{Y}_\tau = E\hat{Y}_\tau \ \ \forall \tau = 1, 2, \dots, N$$

El intervalo de confianza para el predictor del valor esperado se puede calcular como:

$$\hat{Y}_\tau \pm t_{\alpha/2}^{T-K-1} S_{e_\tau} \ \ \forall \tau = 1, 2, \dots, N$$

```
# Intervalo de confianza para un nivel de significación del 5% para la observación 11
? genr LIIPM11 = Cp11 - talpham * sepm11
Se ha generado el escalar LIIPM11 = 12.4384

? genr LSIPM11 = Cp11 + talpham * sepm11
Se ha generado el escalar LSIPM11 = 13.5588

# Intervalo de confianza para un nivel de significación del 5% para la observación 12
? genr LIIPM12 = Cp12 - talpham * sepm12
Se ha generado el escalar LIIPM12 = 11.7721

? genr LSIPM12 = Cp12 + talpham * sepm12
Se ha generado el escalar LSIPM12 = 12.4324
```